# Efficient Steganography with Provable Security Guarantees

Aggelos Kiayias* akiayias@cse.uconn.edu
Yona Raekow yona@cse.uconn.edu
Alexander Russell† acr@cse.uconn.edu
Narasimha Shashidhar karpoor@cse.uconn.edu

Department of Computer Science and Engineering
University of Connecticut, Storrs, CT

**Abstract**

We provide a new provably-secure steganographic encryption protocol that is proven secure in the complexity-theoretic framework of Hopper et al.

The fundamental building block of our steganographic encryption protocol is a "one-time stegosystem" that allows two parties to transmit messages of length shorter than the shared key with *information-theoretic* security guarantees. The employment of a pseudorandom generator (PRG) permits secure transmission of longer messages in the same way that such a generator allows the use of one-time pad encryption for messages longer than the key in symmetric encryption. The advantage of our construction, compared to that of Hopper et al., is that it avoids the use of a pseudorandom function family and instead relies (directly) on a pseudorandom generator in a way that provides linear improvement in the number of applications of the underlying one-way permutation per transmitted bit. This advantageous trade-off is achieved by substituting the pseudorandom function family employed in the previous construction with an appropriate combinatorial construction that has been used extensively in derandomization, namely almost $t$-wise independent function families.

**Keywords:** Information hiding, steganography, data hiding, steganalysis, covert communication.

## 1 Introduction

In a canonical steganographic scenario, Alice and Bob wish to communicate securely in the presence of an adversary, called the "Warden," who monitors whether they exchange "conspicuous" messages. In particular, Alice and Bob may exchange messages that adhere to a certain channel distributions that represents "inconspicuous" communication. By controlling the messages that are transmitted over such a channel, Alice and Bob may exchange messages that cannot be detected by the Warden. There have been two approaches in formalizing this problem, one based on information theory [2, 13, 7] and one based on complexity theory [6]. The latter approach is more concrete and has the potential of allowing more efficient constructions. Most steganographic constructions supported by provable security guarantees are instantiations of the following basic procedure (often referred to as "rejection-sampling").

The problem specifies a family of message distributions (the "channel distributions") that provide a number of possible options for a so-called "covertext" to be transmitted. Additionally, the sender and the receiver possess some sort of private information (typically a keyed hash function, MAC, or other similar function) that maps channel messages to a single bit. In order to send a message bit $m$, the sender draws a

*Supported by NSF CAREER grant CCR-0447808.
†Supported by NSF CAREER grant CCR-0093065, and NSF grants CCR-0121277, CCR-0220264, CCR-0311368, and EIA-0218443.

covertext from the channel distribution, applies the function to the covertext and checks whether it happens to produce the "stegotext" $m$ he originally wished to transmit. If this is the case, the covertext is transmitted. In case of failure, this procedure is repeated. While this is a fairly concrete procedure, there are a number of choices to be made with both practical and theoretical significance. From the security viewpoint, one is primarily interested in the choice of the function that is shared between the sender and the receiver. From a practical viewpoint, one is primarily interested in how the channel is implemented and whether it conforms to the various constraints that are imposed on it by the steganographic protocol specifications (e.g., are independent draws from the channel allowed? does the channel remember previous draws? etc.).

As mentioned above, the security of a stegosystem can be naturally phrased in information-theoretic terms (cf. [2]) or in complexity-theoretic terms [6]. Informally, the latter approach considers the following experiment for the warden-adversary: The adversary selects a message to be embedded and receives either covertexts that embed the message or covertexts simply drawn from the channel distribution (without any embedding). The adversary is then asked to distinguish between the two cases. Clearly, if the probability of success is very close to $1/2$ it is natural to claim that the stegosystem provides security against such (eavesdropping) adversarial activity. Formulation of stronger attacks (such as active attacks) is also possible. Given the above framework, Hopper et al. [6] provided a provably secure stegosystem that pairs rejection sampling with a pseudorandom function family. Given that rejection sampling, when implemented properly and paired with a truly random function, is indistinguishable from the channel distribution, the security of their construction followed from the pseudorandom function family assumption. From the efficiency viewpoint, this construction required about 2 evaluations of the pseudorandom function per bit transmission. Constructing efficient pseudorandom functions is possible either generically [5] or, more efficiently, based on specific number-theoretic assumptions [9]. Nevertheless, pseudorandom function families are a conceptually complex and fairly expensive cryptographic primitive. For example, the evaluation of the Naor-Reingold pseudorandom function on an input $x$ requires $O(|x|)$ modular exponentiations. Similarly, the generic construction [5] requires $O(k)$ PRG doublings of the input string where $k$ is the length of the key.

In this article we take an alternative approach to the design of provably secure stegosystems. Our main contribution is the design of a building block that we call a *one-time stegosystem*: this is a steganographic protocol that is meant to be used for a single message transmission and is proven secure in an information-theoretic sense, provided that the key that is shared between the sender and the receiver is of sufficient length (this length analysis is part of our result). In particular we show that we can securely transmit an $n$ bit message with a key of length $O(n + \log |\Sigma|)$; here $\Sigma$ is the size of the channel alphabet (see Section 3.4 for more details regarding the exact complexity). Our basic building block is a natural analogue of a one time-pad for steganography. It is based on the rejection sampling technique outlined above in combination with an explicit almost $t$-wise independent [1] family of functions. We note that such combinatorial constructions have been extremely useful for derandomization methods and here, to the best of our knowledge, are employed for the first time in the design of steganographic protocols. Given a one-time stegosystem, it is fairly straightforward to construct provably secure steganographic encryption for longer messages by using a pseudorandom generator (PRG) to stretch a random seed that is shared by the sender and the receiver to sufficient length.

The resulting stegosystem is provably secure in the computational sense of Hopper et al. [6] and is in fact much more efficient: in particular, while the Hopper, et al. stegosystem requires 2 evaluations *per bit* of a pseudorandom function, amounting to a linear (in the key-size) number of applications of the underlying PRG (in the standard construction for pseudorandom functions of [5]), in our stegosystem we require *per bit* a constant number of PRG applications.

## 2 Definitions and Tools

We say that a function $\mu : \mathbb{N} \to \mathbb{R}$ is *negligible* if for every positive polynomial $p(\cdot)$ there exists an $N$ such that for all $n > N$, $\mu(n) < \frac{1}{p(n)}$.

We let $\Sigma = \{\sigma_1, \ldots, \sigma_s\}$ denote an alphabet and treat the *channel*, which will be used for data transmission, as a family of random variables $\mathcal{C} = \{C_h\}_{h \in \Sigma^*}$; each $C_h$ is supported on $\Sigma$. These channel distributions

model a history-dependent notion of channel data: if $h_1, h_2, \ldots, h_\ell$ have been sent along the channel thus far, $C_{h_1,\ldots,h_\ell}$ determines the distribution of the next channel element.

**Definition 1.** *A* one-time stegosystem *consists of three probabilistic polynomial time algorithms*

$$S = (SK, SE, SD)$$

*where:*

- *$SK$ is the* key generation algorithm*; we write $SK\,(1^n, \log(1/\epsilon_{sec})) = k$. It takes as input, the security parameter $\epsilon_{sec}$ and the length of the message $n$ and produces a key $k$ of length $\kappa$. (We typically assume that $\kappa = \kappa(n)$ is a monotonically increasing function of $n$.)*
- *$SE$ is the* embedding procedure*, which can access the channel; $SE\,(1^n, k, m, h) = s \in \Sigma^*$. It takes as input the length of the message $n$, the key $k$, a message $m \in M_n \triangleq \{0,1\}^n$ to be embedded, and the history $h$ of previously drawn covertexts. The output is the stegotext $s \in \Sigma^*$.*
- *$SD$ is the* extraction procedure*; $SD\,(1^n, k, c \in \Sigma^*) = m$ or $\mathsf{fail}$. It takes as input $n$, $k$, and some $c \in \Sigma^*$. The output is a message $m$ or the token $\mathsf{fail}$.*

Recall that the *min entropy* of a random variable $X$, taking values in a set $V$, is the quantity

$$H_\infty(X) \triangleq \min_{v \in V} \left(-\log \Pr[X = v]\right) .$$

We say that a channel $\mathcal{C}$ has min entropy $\delta$ if for all $h \in \Sigma^*$, $H_\infty(C_h) \geq \delta$.

**Definition 2** (Soundness)**.** *A stegosystem $(SK, SE, SD)$ is said to be* $(s(\kappa), \delta)$-sound *provided that for all channels $\mathcal{C}$ of minimum entropy $\delta$,*

$$\forall m \in M_n, \Pr[SD(1^n, k, SE(1^\kappa, k, m, h)) \neq m \mid k \leftarrow SK(1^n, \log(1/\epsilon_{sec}))] \leq \ s(\kappa) .$$

One-time stegosystem security is based on the indistinguishability between a transmission that contains a steganographically embedded message and a transmission that contains no embedded messages. An adversary $\mathcal{A}$ against a one-time stegosystem $S = (SK, SE, SD)$ is a pair of algorithms $\mathcal{A} = (SA_1, SA_2)$, that plays the following game, denoted $G^{\mathcal{A}}(1^n)$:

1. A key $k$ is generated by $SK(1^n, \log(1/\epsilon_{sec}))$.
2. Algorithm $SA_1$ receives as input the length of the message $n$ and outputs a triple $(m^*, s, h_{\mathsf{c}}) \in M_n \times \{0,1\}^* * \Sigma^*$, where $s$ is some additional information that will be passed to $SA_2$. $SA_1$ is provided access to $\mathcal{C}$ via an oracle $\mathcal{O}(h)$, which takes the history $h$ as input.

   $\mathcal{O}(\cdot)$, on input $h$, returns to $SA_1$ an element $c$ selected according to $C_h$.
3. A bit $b$ is chosen uniformly at random.
   - If $b = 0$ let $c^* \leftarrow SE(1^n, k, m^*, h)$, so $c^*$ is a stegotext.
   - If $b = 1$ let $c^* = c_1 \circ \cdots \circ c_\lambda$, where $\circ$ denotes string concatenation and $c_i \xleftarrow{r} C_{h \circ \mathsf{c}_1 \circ \cdots \circ \mathsf{c}_{i-1}}$.
4. The input for $SA_2$ is $1^n$, $h_c$, $c^*$ and $s$. $SA_2$ outputs a bit $b'$. If $b' = b$ then we say that $(SA_1, SA_2)$ *succeeded* and write $G^{\mathcal{A}}(1^n) = \text{success}$.

The *advantage* of the adversary $\mathcal{A}$ over a stegosystem $S$ is defined as:

$$\mathbf{Adv}_S^{\mathcal{A}}(n) = \left| \Pr\left[G(1^n) = \text{success}\right] - \frac{1}{2} \right| .$$

The probability includes the coin tosses of $\mathcal{A}$ and $SE$, as well as the coin tosses of $G(1^\kappa)$. The (information-theoretic) insecurity of the stegosystem is defined as

$$\mathbf{InSec}_S(n) = \max_{\mathcal{A}} \{\mathbf{Adv}_S^{\mathcal{A}}(n)\} ,$$

this maximum taken over all (time unbounded) adversaries $\mathcal{A}$.

**Definition 3.** *(Security) We say that a stegosystem is* $(t(n), \delta)$-secure *if for all channels with min entropy* $\delta$ *we have* $\mathbf{InSec}_S(n) \leq t(n)$.

## 2.1 Error-correcting Codes

Our steganographic construction requires an efficient family of codes that can recover from errors introduced by certain binary symmetric channels. In particular, we require an efficient version of the Shannon coding theorem [11, 10]. For an element $x \in \{0,1\}^n$, we let $B_p(x)$ be the random variable equal to $x \oplus e$, where $e \in \{0,1\}^n$ is a random error vector defined by independently assigning each $e_i = 1$ with probability $p$. (Here $x \oplus e$ denotes the vector with $i$th coordinate equal to $x_i \oplus e_i$.)

The classical coding theorem asserts that for every pair of real numbers $0 < R < C \leq 1$ and $n \in \mathbb{N}$, there is a binary code $A_n \subset \{0,1\}^n$, with $\log |A|/n \geq R$, so that for each $a \in A$, maximum-likelihood decoding recovers $a$ from $B_p(a)$ with probability $1 - e^{-\theta(n)}$, where

$$H(p) = p \log p^{-1} + (1-p) \log (1-p)^{-1} = 1 - C \, .$$

The quantity $C$ (determined by $p$), is the *capacity* of the binary symmetric channel induced by $B_p$; the quantity $R = \log |A|/n$ is the *rate* of the code $A$. In this language, the coding theorem asserts that at rates lower than capacity, codes exist that correct random errors with exponentially decaying failure probability.

We formalize our requirements below:

**Definition 4.** *An error-correcting code is a pair of functions* $\mathsf{E} = (Enc, Dec)$*, where* $Enc : \{0,1\}^n \to \{0,1\}^\ell$ *is the* encoding function *and* $Dec : \{0,1\}^\ell \to \{0,1\}^n$ *the corresponding* decoding function*. Specifically, we say that* $\mathsf{E}$ *is a* $(n, \ell, p, \epsilon)$-code *if for all* $m \in \{0,1\}^n$*,*

$$\Pr[Dec(Enc(m) \oplus e) = m] \geq 1 - \epsilon$$

*where* $e = (e_1, \ldots, e_\ell)$ *and each* $e_i$ *is independently distributed in* $\{0,1\}$ *so that* $\Pr[e_i = 1] \leq p$*. We say that* $\mathsf{E}$ *is* efficient *if both Enc and Dec are computable in polynomial time.*

**Proposition 1.** *Let* $\tau = \tau(n)$ *lie in the interval* $(0, 1/4)$*,* $p = 1/2 - \tau$*, and* $R' = 1 - H(p)$*. Let* $n \geq 16$ *be a message length for which* $(104 \log (\log n))^3 / \log n \leq \tau^2$*. Then there is an efficient family of* $(n, \ell(n), p, \epsilon(n))$*-error-correcting codes* $\mathsf{E}_n$ *for which*

$$\epsilon(n) \leq e^{-4n/\log n} \quad \textit{and} \quad \ell(n) \leq (1 + 57/\sqrt[3]{\tau^2 \log n})^2 n/R' \, .$$

*Proof.* This is a consequence of Forney's [3] efficient realizations of the Shannon coding theorem [11, 10]; we work out the technical details in the full version of the paper. □

We refer to [12, 4] for detailed discussions of error-correcting codes over binary symmetric channels.

## 2.2 Function Families and Almost $t$-wise Independence

We will employ the notion of (almost) $t$-wise independent function families (cf. [1], [8]).

**Definition 5.** *A family* $\mathcal{F}$ *of Boolean functions on* $\{0,1\}^n$ *is said to be* $\epsilon$-away from $t$-wise independent *or* $(n, t, \epsilon)$-independent *if for any* $t$ *distinct domain elements* $q_1, q_2, \ldots, q_t$ *we have*

$$\sum_{\alpha \in \{0,1\}^t} \left| \Pr_f [f_k(q_1) f_k(q_2) \cdots f_k(q_t) = \alpha] - \frac{1}{2^t} \right| \leq \epsilon \, , \tag{1}$$

*where* $f$ *chosen uniformly from* $\mathcal{F}$*.*

The above is equivalent to the following formulation quantified over all computationally unbounded adversaries $\mathcal{A}$:

$$\left| \Pr_{f \xleftarrow{r} \mathcal{F}} [G^{\mathcal{A}^{f[t]}}(1^\kappa) = 1] - \Pr_{f \xleftarrow{r} \mathcal{R}} [G^{\mathcal{A}^{f[t]}}(1^\kappa) = 1] \right| \leq \epsilon \,, \tag{2}$$

where $\mathcal{R}$ is the collection of *all* functions from $\{0,1\}^n$ to $\{0,1\}$ and $\mathcal{A}^{f[t]}$ is an unbounded adversary that is allowed to determine up to $t$ queries to the function $f$ before he outputs his bit.

**Lemma 2.** *$\mathcal{F}_\kappa$ is $\epsilon'$-away from $t$-wise independence according to equation (1) if and only if $\mathcal{F}_\kappa$ is $\epsilon'$-away from $t$-wise independence according to equation (2) above.*

We employ the construction of almost $t$-wise independent sample spaces given by [8], [1].

**Theorem 3** ([8], [1])**.** *There exist families of Boolean functions $\mathcal{F}^n_{t,\epsilon}$ on $\{0,1\}^n$ that are $\epsilon$-away from $t$-wise independent, are indexed by keys of length $(2+o(1))(\log\log n + \frac{t}{2} + \log 1/\epsilon)$, and are computable in polynomial time.*

## 2.3 Rejection Sampling

A common method used in steganography employing a channel distribution is that of *rejection sampling* (cf. [2, 6]). Assuming that one wishes to transmit a single bit $m$ and employs a random function $f : \{0,1\}^d \times \Sigma \to \{0,1\}$ that is secret from the adversary, one performs the following "rejection sampling" process:

| $\texttt{rejsam}^f_h(m)$ |
|---|
| $c \xleftarrow{r} C_h$ |
| if $f(c) \neq m$ |
| then $c \xleftarrow{r} C_h$ |
| $\texttt{Output: } c$ |

Here, as above, $\Sigma$ denotes the output alphabet of the channel, $h$ denotes the history of the channel data at the start of the process, and $C_h$ denotes the distribution on $\Sigma$ given by the channel after history $h$. The receiver (also privy to the function $f$) applies the function to the received message $c \in \Sigma$ and recovers $m$ with probability greater than $1/2$. The sender and the receiver may employ a joint state denoted by $i$ in the above process (e.g., a counter), that need not be secret from the adversary. Note that the above process performs only two draws from the channel with the *same* history (more draws could, in principle, be performed). These draws are assumed to be independent. One basic property of rejection sampling that we use is:

**Lemma 4.** *If $f$ is drawn uniformly at random from the collection of all functions $\mathcal{R} = \{f : \Sigma \to \{0,1\}\}$ and $\mathcal{C}$ has min entropy $\delta$, then*

$$\Pr_{f \leftarrow \mathcal{R}} [f(\texttt{rejsam}^f_h(m)) = m] \geq \frac{1}{2} + \tau \ ,$$

*where $\tau = \frac{1}{4}\left(1 - \frac{1}{2^\delta}\right)$.*

*Proof.* Define the event $E$ to be

$$E = [f(\mathsf{c}_1) = m] \vee [f(\mathsf{c}_1) \neq m \wedge f(\mathsf{c}_2) = m] \,;$$

thus $E$ is the event that rejection sampling is successful for $m$. Here $\mathsf{c}_1, \mathsf{c}_2$ are two independent random variables distributed according to the channel distribution $C_h$ and $h$ is determined by the history of channel usage. Recalling that $\Sigma = \{\sigma_1, \ldots, \sigma_s\}$ is the support of the channel distribution $C_h$, let $p_i = \Pr[C_h = \sigma_i]$ denote the probability that $\sigma_i$ occurs. As $f$ is chosen uniformly at random,

$$\Pr[f(\mathsf{c}_1) = m] = 1/2 \,.$$

Then $\Pr[E] = 1/2 + \Pr[A]$, where $A$ is the event that $f(\mathsf{c}_1) \neq m \wedge f(\mathsf{c}_2) = m$. To bound $\Pr[A]$, let $D$ denote the event that $\mathsf{c}_1 \neq \mathsf{c}_2$. Observe that conditioned on $D$, $A$ occurs with probability exactly $1/4$; on the other hand, $A$ cannot occur simultaneously with $\overline{D}$. Thus

$$\Pr[E] = \frac{1}{2} + \Pr[A \mid D] \cdot \Pr[D] + \Pr[A \mid \overline{D}] \cdot \Pr[\overline{D}] = \frac{1}{2} + \frac{1}{4}\Pr[D] \ .$$

To bound $\Pr[D]$, note that

$$\Pr[\bar{D}] = \sum_i p_i^2 \leq \max_i p_i \sum_i p_i = \max_i p_i$$

and hence that $\Pr[D] \geq 1 - \max_i p_i$. Considering that $H_\infty(C) \geq \delta$, we have $\max_i p_i \leq \frac{1}{2^\delta}$ and the success probability is

$$\Pr[E] \geq \frac{1}{2} + \frac{1}{4} \cdot (1 - p_i) \geq \frac{1}{2} + \frac{1}{4}\left(1 - \frac{1}{2^\delta}\right) = \frac{1}{2} + \tau \,,$$

where $\tau = \frac{1}{4}\left(1 - \frac{1}{2^\delta}\right)$. □

# 3 The construction

In this section we outline our construction of a one-time stegosystem as an interaction between Alice (the sender) and Bob (the receiver). Alice and Bob wish to communicate over a channel with distribution $\mathcal{C}$. We assume that $\mathcal{C}$ has min entropy $\delta$, so that $\forall h \in \Sigma^*$, $H_\infty(C_h) \geq \delta$. As above, let $\tau = \frac{1}{4}\left(1 - \frac{1}{2^\delta}\right)$. For simplicity, we assume that the support of $\mathcal{C}_h$ is of size $|\Sigma| = 2^b$.

## 3.1 A one-time stegosystem

Fix an alphabet $\Sigma$ for the channel and choose a message length $n$ and security parameter $\epsilon_{\mathcal{F}}$. Alice and Bob agree on the following:

**An error-correcting code.** Let $\mathsf{E} = (Enc, Dec)$ be an efficient $(n, \lambda, \frac{1}{2} - \tau, \epsilon_{\text{enc}})$-error-correcting code;

**A pseudorandom function family.** Let $\mathcal{F}$ be a function family that is $(\log\lambda + \log|\Sigma|, 2\lambda, \epsilon_{\mathcal{F}})$-independent. We treat elements of $\mathcal{F}$ as Boolean functions on $\{1, \ldots, \lambda\} \times \Sigma$ and, for such a function $f$ we let $f_i : \Sigma \to \{0,1\}$ denote the function $f_i(\sigma) = f(i, \sigma)$.

We will analyze the stegosystem below in terms of arbitrary parameters $\lambda$, $\epsilon_{\mathcal{F}}$, and $\epsilon_{\text{enc}}$, relegating discussion of how these parameters determine the overall efficiency of the system to Section 3.4.

Key generation consists of selecting an element $f \in \mathcal{F}$. Alice and Bob then communicate using the algorithms $SE$ for embedding and $SD$ for extracting as described in Figure 1. In $SE$, after applying the error-correcting code $\mathsf{E}$, we use $\mathtt{rejsam}_h^{f_i}(m_i)$ to obtain an element $c_i$ of the channel for each bit $m_i$ of the message. The resulting stegotext $c_1 \ldots c_\lambda$ is denoted $c_{stego}$. In $SD$ the received stegotext is parsed block by block by evaluating the key function $f_i$ at $c_i$; this results in a message bit. After performing this for each received block, a message of size $\lambda$ is received, which is subjected to decoding via Dec. Note that we sample at most twice from the channel for each bit we wish to send. The error-correcting code is needed to recover from the errors introduced by this process. The detailed security and correctness analysis follow in the next two sections.

## 3.2 Correctness

We focus on the mapping between $\{0,1\}^\lambda$ and $\Sigma^\lambda$ determined by the $SE$ procedure of the one-time stegosystem. In particular, for an initial history $h$ and a key function $f : \{1, \ldots, \lambda\} \times \Sigma \to \{0,1\}$, recall that the covertext of the message $m$ is given by the procedure $P^f(m) = P_h^f(m)$, described in Figure 2; here $h$ is the initial history. We remark now that the procedure defining $P^f$ samples $f$ at no more than $2\lambda$ points and that

| PROCEDURE $SE$: | PROCEDURE $SD$: |
|---|---|
| Input: Key $k$, hidden text $m'$, history $h$ | Input: Key $k$, stegotext $c_{\text{stego}}$ |
| let $m = Enc(m')$ | |
| parse $m$ as $m = m_1 m_2 \dots m_\lambda$ | parse $c_{stego}$ as $c = c_1 c_2 \dots c_\lambda$ |
| for $i = 1$ to $\lambda$ { | for $i = 1$ to $\lambda$ { |
| $c_i = \texttt{rejsam}_h^{f_i}(m_i)$ | set $\bar{m}_i = f_i(c_i)$ |
| set $h \leftarrow h \circ c_i$ | let $\bar{m} = \bar{m}_1 \bar{m}_2 \dots \bar{m}_\lambda$ |
| } | } |
| Output: $c_{stego} = c_1 c_2 \dots c_\lambda \in \Sigma^\lambda$ | Output: $Dec(\bar{m})$ |

Figure 1: Encryption and Decryption algorithms for the one-time stegosystem of 3.1.

$$\mathrm{P}_h^f : \{0,1\}^\lambda \to \Sigma^\lambda$$

input: $h$, $m = m_1 \dots m_\lambda \in \{0,1\}^\lambda$
for $i = 1$ to $\lambda$
$c_i = \texttt{rejsam}_h^{f_i}(m_i)$
$h \leftarrow h \circ c_i$
output: $c = c_1 \dots c_\lambda \in \Sigma^\lambda$

Figure 2: The procedure $P_h^f$.

the family $\mathcal{F}$ used in $SE$ is $\epsilon_{\mathcal{F}}$-away from $2\lambda$-wise independent. For a string $c = c_1 \dots c_\lambda \in \Sigma^\lambda$ and a function $f$, let $R^f(c) = (f_1(c_1), \dots, f_\lambda(c_\lambda)) \in \{0,1\}^\lambda$. If $f$ were chosen uniformly among *all* Boolean functions on $\{1, \dots, \lambda\} \times \Sigma$ then we could conclude from Lemma 4 above that each bit is independently recovered by this process with probability at least $\frac{1}{2} + \tau$. As $\mathsf{E}$ is an $(n, \lambda, \frac{1}{2} - \tau, \epsilon_{\text{enc}})$-error-correcting code, this would imply that

$$\Pr_{f \leftarrow \mathcal{R}}[R^f(P_h^f(m)) = m] \geq 1 - \epsilon_{\text{enc}} \ .$$

This is a restatement of the correctness analysis of Hopper, et al [6]. Recalling that the procedure defining $R^f(P_h^f(\cdot))$ involves no more than $2\lambda$ samples of $f$, condition (2) following Definition 5 implies that

$$\Pr_{f \leftarrow \mathcal{F}}[R^f(P_h^f(m)) = m] \geq 1 - \epsilon_{\text{enc}} - \epsilon_{\mathcal{F}} \tag{3}$$

so long as $\mathcal{F}$ is $(\log \lambda + \log |\Sigma|, 2\lambda, \epsilon_{\mathcal{F}})$-independent. (We remark that as described above, the procedure $P_h^f$ depends on the behavior of channel; note, however, that if there were a sequence of channel distributions which violated (3) then there would be a fixed sequence of channel responses, and thus a deterministic process $P^f$, which also violated (3).) To summarize

**Lemma 5.** *With $SE$ and $SD$ described as above, the probability that a message $m$ is recovered from the stegosystem is at least $1 - \epsilon_{enc} - \epsilon_{\mathcal{F}}$.*

### 3.3 Security

In this section we argue about the security of our one-time stegosystem. First we will observe that the output of the rejection sampling function $\texttt{rejsam}_h^f$, with a truly random function $f$, is indistinguishable from the channel distribution $\mathcal{C}_h$. (This is a folklore result implicit in previous work.) We then show that if $f$ is selected from a family that is $\epsilon_{\mathcal{F}}$-away from $2\lambda$-wise independent, the advantage of an adversary $\mathcal{A}$ to distinguish between the output of the protocol and $\mathcal{C}_h$ is bounded above by $\epsilon_{\mathcal{F}}$. Let $\mathcal{R} = \{f : \Sigma \to \{0,1\}\}$. First we characterize the probability distribution of the rejection sampling function:

**Proposition 6.** *The function* $\texttt{rejsam}_h^f(m)$ *is a random variable with probability distribution expressed by the following function: Let* $c \in \Sigma$ *and* $m \in \{0,1\}$*. Let* $\mathit{miss}_f(m) = \Pr_{c' \leftarrow \mathcal{C}_h}[f(c') \neq m]$ *and* $p_c = \Pr_{c' \leftarrow \mathcal{C}_h}[c' = c]$*. Then*

$$\Pr[\texttt{rejsam}_h^f(m) = c] = \begin{cases} p_c \cdot (1 + \mathit{miss}_f(m)) & \text{if } f(c) = m \ , \\ p_c \cdot \mathit{miss}_f(m) & \text{if } f(c) \neq m \ . \end{cases}$$

*Proof.* Let $c_1$ and $c_2$ be the two (independent) samples drawn from $\mathcal{C}_h$ during rejection sampling. (For simplicity, we treat the process as having drawn two samples even in the case where it succeeds on the first draw.) Note, now, that in the case where $f(c) \neq m$, the value $c$ is the result of the rejection sampling process precisely when $f(c_1) \neq m$ and $c_2 = c$; as these samples are independent, this occurs with probability $\mathsf{miss}_f(m) \cdot p_c$.

In the case where $f(c) = m$, however, we observe $c$ whenever $c_1 = c$ or $f(c_1) \neq m$ and $c_2 = c$. As these events are disjoint, their union occurs with probability $p_c \cdot (\mathsf{miss}_f(m) + 1)$, as desired. □

**Lemma 7.** *For any* $h \in \Sigma^*, m \in \{0,1\}$*, the random variable* $\texttt{rejsam}_h^f(m)$ *is perfectly indistinguishable from the channel distribution* $C_h$ *when* $f$ *is drawn uniformly at random from the space of* $\mathcal{R}$*.*

*Proof.* Let $f$ be a random function, as described in the statement of the lemma. Fixing the elements $c$, and $m$, we condition on the event $E_{\neq}$, that $f(c) \neq m$. In light of Proposition 6, for any $f$ drawn under this conditioning we shall have that $\Pr[\texttt{rejsam}_h^f(m) = c]$ is equal to

$$\Pr_{c' \leftarrow \mathcal{C}_h}[c' = c] \cdot \mathsf{miss}_f(m) = p_c \cdot \mathsf{miss}_f(m) \ ,$$

where we have written $\mathsf{miss}_f(m) = \Pr_{c' \leftarrow \mathcal{C}_h}[f(c') \neq m]$ and $p_c = \Pr_{c' \leftarrow \mathcal{C}_h}[c' = c]$. Conditioned on $E_{\neq}$, then, the probability of observing $c$ is

$$\mathbf{E}_f[p_c \cdot \mathsf{miss}_f(m) \mid E_{\neq}] = p_c \left( p_c + \frac{1}{2}(1 - p_c) \right) \ .$$

Letting $E_{=}$ be the event that $f(i, c) = m$, we similarly compute

$$\mathbf{E}_f[p_c \cdot \mathsf{miss}_f(m) \mid E_{=}] = p_c \left( 1 + \frac{1}{2}(1 - p_c) \right) \ .$$

As $\Pr[E_{=}] = \Pr[E_{\neq}] = 1/2$, we conclude that the probability of observing $c$ is exactly

$$\frac{1}{2} \left( p_c \left( p_c + \frac{1 - p_c}{2} \right) + p_c \left( 1 + \frac{1 - p_c}{2} \right) \right) = p_c \ ,$$

as desired. □

The following corollary follows immediately from the lemma above.

**Corollary 8.** *For any* $h \in \Sigma^*, m \in \{0,1\}^\lambda$*, the random variable* $P_h^f$ *is perfectly indistinguishable from the channel distribution* $C_h^\lambda$ *when* $f$ *is drawn uniformly at random from the space of all Boolean functions on* $\{1, \ldots, \lambda\} \times \Sigma$*.*

Having established the behavior of the rejection sampling function when a truly random function is used, we proceed to examine the behavior of rejection sampling in our setting where the function is drawn from a function family that is $\epsilon_{\mathcal{F}}$-away from $2\lambda$-wise independence. In particular we will show that the insecurity of the defined stegosystem is characterized as follows:

**Lemma 9.** *The insecurity of the stegosystem* $S$ *of Section 3.1 is bound by* $\epsilon_{\mathcal{F}}$*, i.e.,* $\mathbf{InSec}_S(n) \leq \epsilon_{\mathcal{F}}$*, where* $\epsilon_{\mathcal{F}}$ *is the bias of the almost* $2\lambda$*-wise independent function family employed; recall that* $\lambda = \ell(n)$ *is the stretching of the input incurred due to the error-correcting code.*

*Proof.* Let us play the following game $G(1^\kappa)$ with the adversary $\mathcal{A}$.

In each round we either select $G_1^{\mathcal{A}}$ or $G_2^{\mathcal{A}}$:

| $G_1^{\mathcal{A}}(1^\kappa)$ | |
|---|---|
| 1. | $k \leftarrow \{0,1\}^\kappa$ |
| 2. | $(m^*, s) \leftarrow SA_1^{\mathcal{O}(h)}(1^\kappa, h)$, $m^* \in \{0,1\}^n$ |
| 3. | $b \xleftarrow{r} \{0,1\}$ |
| 4. | $c^* = \begin{cases} c_0, c_1, \ldots c_{\lambda-1} & c_i = \mathtt{rejsam}_h^{f_\kappa, i}(m_i), h = h \circ c \quad \text{if } b = 0 \\ \text{from the channel} & \text{if } b = 1 \end{cases}$ |
| 5. | $b^* \leftarrow SA_2(c^*, s)$ |
| 6. | if $b = b^*$ then success |

| $G_2^{\mathcal{A}}(1^\kappa)$ | |
|---|---|
| 1. | $f \leftarrow \mathcal{R}$ |
| 2. | $(m^*, s) \leftarrow SA_1^{\mathcal{O}(h)}(1^\kappa, h)$, $m^* \in \{0,1\}^n$ |
| 3. | $b \xleftarrow{r} \{0,1\}$ |
| 4. | $c^* = \begin{cases} c_0, c_1, \ldots c_{\lambda-1} & c_i = \mathtt{rejsam}_h^{f, i}(m_i), h = h \circ c \quad \text{if } b = 0 \\ \text{from the channel} & \text{if } b = 1 \end{cases}$ |
| 5. | $b^* \leftarrow SA_2(c^*, s)$ |
| 6. | if $b = b^*$ then success |

$$\begin{aligned} \mathbf{Adv}_S^{\mathcal{A}}(G(1^\kappa)) &= \left| \Pr[\mathcal{A}^{\mathcal{O}(h), c^* \leftarrow SE(k,\cdot,\cdot,\cdot)} = 1] - \Pr[A^{\mathcal{O}(h), c^* \leftarrow \mathcal{C}_h} = 1] \right| \\ &= \Pr_{f \leftarrow \mathcal{F}_\kappa}[G(1^\kappa) = 1] - \Pr_{f \leftarrow \mathcal{R}}[G(1^\kappa) = 1] \leq \epsilon_{\mathcal{F}} \end{aligned}$$

and the lemma follows by the definition of insecurity. □

### 3.4 Putting it all together

The objective of this section is to integrate the results of the previous sections of the paper into one unifying theorem. As our system is built over two-sample rejection sampling, a process that faithfully transmits each bit with probability $1/2 + \tau$, we cannot hope to achieve rate exceeding

$$R' = 1 - H(1/2 + \tau) = 1 - H(1/4 + 2^{-\delta}/4)\,.$$

Indeed, as described in the theorem below, the system asymptotically converges to the rate of this underlying rejection sampling channel. (We remark that with sufficiently large channel entropy, one can draw more samples during rejection sampling without interfering with security; this can control the noise introduced by rejection sampling.)

**Theorem 10.** *For $\delta = \Omega(\sqrt{(\log\log n)^3/\log n})$ the stegosystem $S$ uses private keys $k$ of length no more than*

$$(2 + o(1))\,[\lambda(n) + \log 1/\epsilon_{\mathcal{F}} + \log\log\log|\Sigma|]$$

*and is both $(\epsilon_{enc} + \epsilon_{\mathcal{F}}, \delta)$-sound and $(\epsilon_{\mathcal{F}}, \delta)$-secure. The length of the stegotext $\lambda(n)$ is*

$$\lambda(n) \leq \left(1 + \frac{1}{\log(\log n)}\right)^2 \frac{n}{R'}\,,$$

*where $\epsilon_{enc} \leq e^{-4n/\log n}$ and $R' = 1 - H\left(1/4 + 2^{-\delta}/4\right)$.*

*Proof.* Let $\Sigma = \{\sigma_1, \ldots, \sigma_s\}$ denote an alphabet and define the *channel* as a family of random variables $\mathcal{C} = \{C_h\}_{h \in \Sigma^*}$; each $C_h$ supported on $\Sigma$. Also, the channel $\mathcal{C}$ has min entropy $\delta$, so that $\forall h \in \Sigma^*$, $H_\infty(C_h) \geq \delta$. Fix an alphabet $\Sigma$ for the channel and choose a message length $n \geq 16$ such that

$$-\log\left(1 - 4\sqrt{104\log(\log n)^3/\log n}\right) \leq \delta\,.$$

Under the assumption that the channel $\mathcal{C}$ has min entropy $\delta$, the binary symmetric channel induced by the rejection sampling process of Lemma 4 has transition probability no more than $1/4(1+2^{-\delta})$. We have an efficient $(n, \lambda, \frac{1}{4}(1+2^{-\delta}), \epsilon_{\text{enc}})$ error-correcting code as discussed in Section 2.1 that encodes messages of length $n$ as codewords of length

$$\lambda(n) = \left(1 + \frac{57}{\sqrt[3]{\tau^2 \log n}}\right)^2 \frac{n}{R'} \leq \left(1 + \frac{1}{\log(\log n)}\right)^2 \frac{n}{R'} \text{ bits}$$

$$(2 + o(1)) \left[\lambda(n) + \log 1/\epsilon_{\mathcal{F}} + \log\log\log |\Sigma|\right]$$

random bits; these serve as the key for the stegosystem. In light of the conclusions of Lemma 9 and Lemma 5, this system achieves the $(\epsilon_{\text{enc}} + \epsilon_{\mathcal{F}}, \delta)$-soundness and $(\epsilon_{\mathcal{F}}, \delta)$-security. □

For concreteness, we record two corollaries:

**Corollary 11.** *There exists a function $\delta(n) = o(1)$ so that the stego system $S$, using private keys $k$ of length no more than*

$$O\left(n + \log|\Sigma| + \log 1/\epsilon_{\mathcal{F}}\right),$$

*is both $(e^{-4n/\log n} + \epsilon_{\mathcal{F}}, \delta)$-sound and $(\epsilon_{\mathcal{F}}, \delta)$-secure. Here, the length of the stegotext is*

$$\lambda(n) = (1 + o(1)) \frac{n}{R'}$$

*where $R' = 1 - H\left(1/4(1+2^{-\delta})\right)$.*

**Corollary 12.** *For any constant $\delta$, the stegosystem $S$ uses private keys of length $O(n + \log\Sigma + \log\epsilon_{\mathcal{F}})$ and transmits no more than $O(n)$ symbols.*

# 4 A provably secure stegosystem for longer messages

In this section we show how to apply the "one-time" stegosystem of Section 3.1 together with a pseudorandom number generator so that longer messages can be transmitted.

**Definition 6.** *Let $U_l$ denote the uniform distribution over $\{0,1\}^l$. A polynomial time deterministic program $G$ is a pseudorandom generator (PRG) if the following conditions are satisfied:*

**Variable output** *For all seeds $x \in \{0,1\}^*$ and $y \in \mathbb{N}$, $|G(x, 1^y)| = y$ and, furthermore, $G(x, 1^y)$ is a prefix of $G(x, 1^{y+1})$.*

**Pseudorandomness** *For every polynomial $p$ the set of random variables $\{G(U_l, 1^{p(l)})\}_{l\in\mathbb{N}}$ is computationally indistinguishable from the uniform distribution $U_{p(l)}$.*

Note that there is a procedure $G'$ that if $z = G(x, 1^y)$ it holds that $G(x, 1^{y+y'}) = G'(x, z, 1^{y'})$ (i.e., if one maintains $z$, one can extract the $y'$ bits that follow the first $y$ bits without starting from the beginning). For a PRG $G$, if $A$ is some statistical test, then we define the advantage of $A$ over the PRNG as follows:

$$\mathbf{Adv}_G^A(l) = \left| \Pr_{\hat{l} \leftarrow G(U_l, 1^{p(l)})} [A(\hat{l}) = 1] - \Pr_{\hat{l} \leftarrow U_{p(l)}} [A(\hat{l}) = 1] \right|$$

The insecurity of the PRNG $G$ is then defined

$$\mathbf{InSec}_G^{PRG}(l) = \max_A \{\mathbf{Adv}_G^A(l)\} \ .$$

Note that typically in PRGs there is a procedure $G'$ as well as the process $G(x, 1^y)$ produces some auxiliary data $\mathsf{aux}_y$ of small length so that the rightmost $y'$ bits of $G(x, 1^{y+y'})$ may be sampled directly as

$G'(x, 1^{y'}, \mathsf{aux}_y)$. Consider now the following stegosystem $S' = (SE', SD')$ that can be used for arbitrary many and long messages and employs a PRG $G$ and the one-time stegosystem $(SK, SE, SD)$ of Section 3.1. The two players Alice and Bob, share a key of length $l$ denoted by $x$. They also maintain a state $N$ that holds the number of bits that have been transmitted already as well the auxiliary information $\mathsf{aux}_N$ (initially empty). The function $SE'$ is given input $N, \mathsf{aux}_N, x, m \in \{0,1\}^n$ where $m$ is the message to be transmitted. $SE'$ in turn employs the PRG $G$ to extract a number of bits $\kappa$ as follows $k = G'(x, 1^\kappa, \mathsf{aux}_N)$. The length $\kappa$ is selected to match the number of key bits that are required to transmit the message $m$ using the one-time stegosystem of section 3.1. Once the key $k$ is produced by the PRG the procedure $SE'$ invokes the one-time stegosystem on input $k, m, h$. After the transmission is completed the history $h$, the count $N$, as well as the auxiliary PRG information $\mathsf{aux}_N$ are updated accordingly. The function $SD'$ is defined in a straightforward way based on $SD$.

**Theorem 13.** *The stegosystem $S' = (SE', SD')$ is provably secure in the model of [6] (universally steganographically secret against chosen hiddentext attacks); in particular*

$$\mathbf{InSec}^{SS}_{S'}(t, q, l) \leq \mathbf{InSec}^{PRG}(t + \gamma(\ell(l)), \ell(l) + \mathsf{polylog}(l))$$

*(where $t$ is the time required by the adversary, $q$ is the number of chosen hiddentext queries it makes, $l$ is the total number of bits across all queries and $\gamma(v)$ is the time required to simulate the $SE'$ oracle for $v$ bits).*

### 4.1 Performance Comparison of the Stegosystem $S'$ and the Hopper, Langford, von Ahn System

The system of Hopper, et al. [6] concerns a situation where the min entropy of all $\mathcal{C}_h$ is at least 1 bit. In this case, we may select an $(n, \lambda, 3/8, \epsilon_{\text{enc}})$-error-correcting code $\mathsf{E}$. Then the system of Hopper, et al. correctly decodes a given message with probability at least $1 - \epsilon_{\text{enc}}$ and makes no more than $2\lambda$ calls to a pseudorandom function family. Were one to use the pseudorandom function family of Goldreich, Goldwasser, and Micali [5], then this involves production of $\Theta(\lambda \cdot k \cdot (\log(|\Sigma|) + \log \lambda))$ pseudorandom bits, where $k$ is the security parameter of the pseudorandom function family. Of course, the security of the system depends on the security of the underlying pseudorandom generator. On the other hand, with the same error-correcting code, the steganographic system described above utilizes $O\left[\log\log\log|\Sigma| + \lambda + \log 1/\epsilon_{\mathcal{F}}\right]$ pseudorandom bits, correctly decodes a given message with probability $1 - (\epsilon_{\text{enc}} + \epsilon_{\mathcal{F}})$, and possesses insecurity no more than $\epsilon_{\mathcal{F}}$. In order to compare the two schemes, note that by selecting $\epsilon_{\mathcal{F}} = 2^{-k}$, both the decoding error and the security of the two systems differ by at most $2^{-k}$, a negligible function in terms of the security parameter $k$. (Note also that pseudorandom functions utilized in the above scheme have security no better than $2^{-k}$ with security parameter $k$.) In this case, the number of pseudorandom bits used by our system,

$$(2 + o(1))\left[\lambda(n) + \log 1/\epsilon_{\mathcal{F}} + \log\log\log|\Sigma|\right],$$

is a dramatic improvement over the $\Theta(\lambda k \log(|\Sigma|\lambda))$ bits of the scheme above.

## References


[1] Noga Alon, Oded Goldreich, Johan Håstad, and René Peralta. Simple construction of almost k-wise independent random variables. *Random Struct. Algorithms*, 3(3):289–304, 1992.

[2] Christian Cachin. An information-theoretic model for steganography. In *Information Hiding*, pages 306–318, 1998.

[3] G. D. Forney, Jr. *Concatenated Codes*. Research Monograph No. 37. MIT Press, 1966.

[4] R. G. Gallager. A simple derivation of the coding theorem and some applications. *IEEE Transactions on Information Theory*, IT-11:3–18, Jan. 1965.

[5] Oded Goldreich, Shafi Goldwasser, and Silvio Micali. How to construct random functions. *J. ACM*, 33(4):792–807, 1986.

[6] Nicholas J. Hopper, John Langford, and Luis von Ahn. Provably secure steganography. In *CRYPTO*, pages 77–92, 2002.

[7] Thomas Mittelholzer. An information-theoretic approach to steganography and watermarking. In *Information Hiding*, pages 1–16, 1999.

[8] Joseph Naor and Moni Naor. Small-bias probability spaces: Efficient constructions and applications. *SIAM J. Comput.*, 22(4):838–856, 1993.

[9] Moni Naor and Omer Reingold. Number-theoretic constructions of efficient pseudo-random functions. *J. ACM*, 51(2):231–262, 2004.

[10] C. E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423 and 623–656, July and October, 1948.

[11] C. E. Shannon and W. Weaver. *The Mathematical Theory of Communication*. University of Illinois Press, Urbana, Illinois, 1949.

[12] J. H. van Lint. *Introduction to Coding Theory*. Number 86 in Graduate Texts in Mathematics. Springer-Verlag, 3rd edition edition, 1998.

[13] Jan Zöllner, Hannes Federrath, Herbert Klimant, Andreas Pfitzmann, Rudi Piotraschke, Andreas Westfeld, Guntram Wicke, and Gritta Wolf. Modeling the security of steganographic systems. In *Information Hiding*, pages 344–354, 1998.